

Yin, S.-C., Kenny, P. and R. Rose. "Experiments in Speaker Adaptation for Factor Analysis Based Speaker Verification" In *Proceedings of IEEE Odyssey 2006 - The Speaker and Language Recognition Workshop*. San Juan, Puerto Rico, June 28-30, 2006.

Abstract

This paper presents methods for supervised and unsupervised speaker adaptation of Gaussian mixture speaker models in text independent speaker verification. The methods are based on an approach which is able to decompose speaker and channel variability so that progressive updating of speaker models can be performed while minimizing the influence of the channel variability associated with the adaptation utterances. This approach relies on joint factor analysis model of intrinsic speaker variability and session variability where inter-session variation is assumed to result primarily from the effects of the channel [1]. These adaptation methods have been evaluated under the adaptation paradigm defined under the NIST 2005 speaker recognition evaluation plan which is based on conversational telephone speech [2]. It was found that when both target speaker model training and speaker verification trials were performed using a five minute excerpt from single conversation, an equal error rate (EER) of 4.5% and minimum detection cost function (DCF) of 0.013 were obtained when performing unsupervised speaker adaptation during evaluation. It will be shown that this performance is comparable to that obtained by state of the art speaker verification systems that rely on a larger set of features and are trained from as many as eight conversations from the target speaker.

¹ P. Kenny, G. Boulianne, P. Ouellet, and P. Dumouchel, "Joint factor analysis versus eigenchannels in speaker recognition," submitted to *IEEE Trans. Audio, Speech, and Language Processing*. [Online]. Available: <http://www.crim.ca/perso/patrick.kenny/>

² (2005) The NIST year 2005 speaker recognition evaluation plan. [Online]. Available: <http://www.itl.nist.gov/iad/894.01/tests/spk/2005>